

IN THE CLAIMS

Please amend the following claims:

1. (Currently Amended) A method for document classification comprising:
using a first directory structure mirroring a second directory structure used by a user for storing documents;
analyzing content of the documents within the second directory structure to determine a plurality of document classes within the second directory structure, the plurality of document classes indicating a user approach to placing documents in the second directory structure;
determining a document classification profile associated with the first directory structure based on the plurality of document classes;
analyzing content of a previously unclassified electronic document to determine a textual profile and a graphical profile of the electronic document;
generating a classification of the document based on the textual profile and the graphical profile; and
storing the electronic document in one or more directories within the first directory structure based on the classification of the document and the document classification profile associated with the first directory structure, to resemble the user approach to placing the documents in the second directory structure.
2. (Previously Presented) The method defined in Claim 1 wherein the first directory structure comprises a hierarchy of documents mirroring in a similar fashion an

organization in the second directory structure representing a pre-existing memory storing documents.

3. (Previously Presented) The method defined in Claim 2 wherein the pre-existing memory comprises a hard disk.

4. (Previously Presented) The method of claim 1, wherein analyzing content of an electronic document comprises:

determining characteristic words of the document;

determining a frequency for each characteristic word; and

building a frequency table based on the frequency associated with each characteristic word.

5. (Previously Presented) The method of claim 1, wherein analyzing content of an electronic document comprises:

determining a point set corresponding to the electronic document, wherein points of the point set correspond to end points of lines;

determining a density of points within the point set;

generating the graphical profile based, at least in part, on the density of points within the point set.

6. (Previously Presented) The method of claim 1, wherein generating a classification of the document based on the textual profile and the graphical profile comprises combining results from textual and graphical analysis using a Borda Count.

7. (Previously Presented) The method defined in Claim 1 further comprising building the first directory structure by building a hierarchy of documents based on organization of documents stored in a hard drive.

8. (Previously Presented) The method defined in Claim 1 further comprising building the first directory structure by extracting graphical and text features from documents in a directory-based memory to obtain a document classification profile of each subdirectory in the directory-based memory.

9. (Currently Amended) A software product including a machine-readable medium having stored thereon sequences of instructions, which, when executed by a processor, cause the processor to:

use a first directory structure mirroring a second directory structure used by a user for storing documents;

analyze content of the documents within the second directory structure to determine a plurality of document classes within the second directory structure, the plurality of document classes indicating a user approach to placing documents in the second directory structure;

determine a document classification profile associated with the first directory structure based on the plurality of document classes;

analyze content of a previously unclassified electronic document to determine a textual profile and a graphical profile of the electronic document;

generate a classification of the document based on the textual profile and the graphical profile; and

store the electronic document in one or more directories within the first directory structure based on the classification of the document and the document classification profile associated with the first directory structure, to resemble the user approach to placing the documents in the second directory structure.

10. (Previously Presented) The machine-readable medium of claim 9, wherein the sequences of instructions that cause the processor to analyze content of an electronic document further comprise sequences of instructions that cause the processor to:

- determine characteristic words of the document;
- determine a frequency for each characteristic word; and
- build a frequency table based on the frequency associated with each characteristic word.

11. (Previously Presented) The machine-readable medium of claim 9, wherein the sequences of instructions that cause the processor to analyze content of an electronic document further comprise sequences of instructions that cause the processor to:

- determine a point set corresponding to the electronic document, wherein points of the point set correspond to end points of lines;
- determine a density of points within the point set;
- generate the graphical profile based, at least in part, on the density of points within the point set.

12. (Previously Presented) The machine-readable medium of claim 9, wherein the sequences of instructions that cause the processor to generate a classification of the document based on the textual profile and the graphical profile further comprises sequences of instructions that cause the processor to combine results from textual and graphical analysis using a Borda Count.

13. (Currently Amended) A method for document classification comprising:
analyzing content of documents within a pre-existing directory structure to determine a plurality of document classes within the pre-existing directory structure, the plurality of document classes indicating a user approach to placing documents in the pre-existing directory structure;

determining a document classification profile of the pre-existing directory structure based on the plurality of document classes;

generating a mirror directory structure based on the pre-existing document directory structure;

receiving a previously unclassified electronic document;

analyzing content of the electronic document to determine a textual profile and a graphical profile of the electronic document; and

placing the electronic document at a certain storage location in the mirror directory structure based on the document classification profile of the pre-existing document directory structure, the textual profile of the document, and the graphical profile of the document, to resemble [[a]] the user approach to placing the documents in the pre-existing directory structure.

14. (Previously Presented) The method of claim 13, wherein analyzing content of documents within the pre-existing document directory structure further comprises:

recursively descending the pre-existing document directory structure;

generating a list of directories in the pre-existing document directory structure;

and

examining files in the directories of the pre-existing document directory structure.

15. (Original) The method of claim 13 wherein the pre-existing document directory structure is a hard disk directory structure.

16. (Original) The method of claim 13 wherein generating a mirror directory structure based on the pre-existing document directory structure comprises generating a document directory structure having a set of directories and relationships equivalent to the pre-existing document directory structure.

17. (Currently Amended) The method of claim 13 wherein placing the electronic document in the mirror directory structure comprises:

determining a primary directory in the pre-existing document directory structure in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

storing the document in a directory [[in]] of the mirror directory structure that corresponds to the primary directory in the pre-existing document directory structure.

18. (Previously Presented) The method of claim 17 further comprising:

determining a secondary directory in the pre-existing document directory in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

storing the document in a directory of the mirror directory structure that corresponds to the secondary directory in the pre-existing document directory structure.

19. (Currently Amended) A computer-readable medium having stored thereon sequences of instructions which, when executed by a processor, cause the processor to:

analyze content of documents within a pre-existing directory structure to determine a plurality of document classes within the pre-existing directory structure, the plurality of document classes indicating a user approach to placing documents in the pre-existing directory structure;

determine a document classification profile of the pre-existing directory structure based on the plurality of document classes;

generate a mirror directory structure based on the pre-existing document directory structure;

receive a previously unclassified electronic document;

analyze content of the electronic document to determine a textual profile and a graphical profile of the electronic document; and

place the electronic document at a certain storage location in the mirror directory structure based on the document classification profile of the pre-existing document directory structure, the textual profile of the document, and the graphical profile of the document, to resemble [[a]] the user approach to placing the documents in the pre-existing directory structure.

20. (Previously Presented) The computer-readable medium of claim 19, wherein the sequences of instructions that cause the processor to analyze content of documents within a pre-existing document directory structure further comprise sequences of instructions that cause the processor to:

recursively descending the pre-existing document directory structure;
generating a list of directories in the pre-existing document directory structure;
and
examining files in the directories of the pre-existing document directory structure.

21. (Previously Presented) The computer-readable medium of claim 19, wherein the sequences of instructions that cause the processor to generate a mirror directory structure further comprise sequences of instructions that cause the processor to generate a document directory structure having a set of directories and relationships equivalent to the pre-existing document directory structure.

22. (Previously Presented) The computer-readable medium of claim 19, wherein the sequences of instructions that cause the processor to place a document in the mirror directory structure further comprise sequences of instructions that cause the processor to:

determine a primary directory in the pre-existing document directory structure in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

store the document in a directory of the mirror directory structure that corresponds to the primary directory in the pre-existing document directory structure.

23. (Previously Presented) The computer-readable medium of claim 22 further comprising sequences of instructions that cause the processor to:

determine a secondary directory in the pre-existing document directory in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

store the document in a directory of the mirror directory structure that corresponds to the secondary directory in the pre-existing document directory structure.

24. (Currently Amended) An apparatus comprising:

means for analyzing content of documents within a pre-existing directory structure to determine a plurality of document classes within the pre-existing directory structure, the plurality of document classes indicating a user approach to placing documents in the pre-existing directory structure;

means for determining a document classification profile of the pre-existing directory structure based on the plurality of document classes;

means for generating a mirror directory structure based on the pre-existing document directory structure;

means for receiving a previously unclassified electronic document;

means for analyzing content of the electronic document to determine a textual profile and a graphical profile of the electronic document; and

means for placing the electronic document at a certain storage location in the mirror directory structure based on the document classification profile of the pre-existing document directory structure, the textual profile of the document, and the graphical profile of the document, to resemble [[a]] the user approach to placing the documents in the pre-existing directory structure.

25. (Previously Presented) The apparatus of claim 24, wherein means for analyzing content of documents within the pre-existing document directory structure further comprises:

means for recursively descending the pre-existing document directory structure;

means for generating a list of directories in the pre-existing document directory structure; and

means for examining files in directories of the pre-existing document directory structure.

26. (Previously Presented) The apparatus of claim 24, wherein means for generating a mirror directory structure comprises means for generating a document directory structure having a set of directories and relationships equivalent to the pre-existing document directory structure.

27. (Previously Presented) The apparatus of claim 24, wherein means for placing a document in the mirror directory structure comprises:

means for determining a primary directory in the pre-existing document directory structure in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

means for storing the document in a directory of the mirror directory structure that corresponds to the primary directory in the pre-existing document directory structure.

28. (Previously Presented) The apparatus of claim 27 further comprising:

means for determining a secondary directory in the pre-existing document directory in which the document is to be placed based on the document classification profile of the pre-existing document directory structure; and

means for storing the document in a directory of the mirror directory structure that corresponds to the secondary directory in the pre-existing document directory structure.

29. (Currently Amended) A document processing system comprising:

a document scanning device;

a document storage device coupled to the document scanning device, wherein the document storage device has a pre-existing document directory structure and a mirror document directory structure organized based on the pre-existing document directory structure; and

a processor coupled to the document scanning device and to the document storage device, wherein the processor is to analyze content of documents within the pre-existing document directory structure to determine a plurality of document classes in the pre-existing document directory structure, the plurality of document classes indicating a user approach to placing documents in the pre-existing directory structure, to determine a

document classification profile of the pre-existing document directory structure based on the plurality of document classes, to analyze content of a document scanned by the document scanning device, to determine which directory in the mirror document directory structure the scanned document is to be placed based on the analysis of the content of the scanned document and the document classification profile of the pre-existing document directory structure, and to store the scanned document in the determined directory in the mirror document directory structure to resemble [[a]] the user approach to placing the documents in the pre-existing directory structure.

30. (Previously Presented) The document processing system of claim 29 wherein the processor is further to determine a secondary directory in the pre-existing document directory structure in which the document is to be placed and to store the document in a corresponding secondary directory in the mirror document directory structure.

31. Cancelled.

32. (Previously Presented) The document processing system of claim 29 wherein the processor is to analyze content of a document scanned by the document scanning device based on image and textual content of the scanned document.